KTH Mathematics

SF2822 Applied nonlinear optimization, final exam
Wednesday June 9 2010 8.00–13.00
Brief solutions

---

1. (a) The first-order necessary optimality conditions for (EQP) are given by $Hx+c=0$. As H is nonsingular, there is a unique solution given by $x^1 = (1\ 1\ 1)^T$.

The matrix H is not positive semidefinite, since the leading two-by-two principal submatrix is indefinite. With $d = (1\ \ -1\ 0)^T$, we obtain $d^T H d = -2$. Consequently, x^1 does not satisfy the second-order necessary optimality conditions to (EQP).

Consequently, there is no point that satisfies the second-order necessary optimality conditions for (EQP).

(b) The first-order necessary optimality conditions for (EQP) are given by

$$\begin{pmatrix} H & A^T \\ A & 0 \end{pmatrix} \begin{pmatrix} x \\ -\lambda \end{pmatrix} = \begin{pmatrix} -c \\ b \end{pmatrix}$$

which has unique solution $x^2 = (0\ 3\ 1)^T$, $\lambda^2 = 3$. We may for example form a matrix Z whose columns form a basis for $\text{null}(A)$ as

$$Z = \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix},$$

for which $Z^T H Z = I$. Hence, x^2 satisfies the second-order necessary optimality conditions.

(c) Since A has only one row, a local minimizer to (IQP) has to be a local minimizer to (QP) or a local minimizer to (EQP). Since x^1 does not satisfy the second-order necessary optimality conditions to (QP), it is not a local mininimizer to (QP). Hence, it is not a local minimizer to (IQP). Since x^2 satisfies the second-order sufficient optimality conditions to (EQP), it is a local minimizer to (EQP). In addition, since $\lambda^2 > 0$, it is also a local minimizer to (IQP).

(d) Let $q(x) = \frac{1}{2}x^T H x + c^T x$. With d given as in (1a), it follows that $q(x^1 + \alpha d)$ and $q(x^1 - \alpha d)$ tend to minus infinity as $\alpha \to \infty$. Since we have only one constraint, at least one of $x^1 + \alpha d$ and $x^1 - \alpha d$ must remain feasible in (IQP) as $\alpha \to \infty$. We conclude that no global minimizer can exist.

2. No constraints are active at the initial point. Hence, the working set is empty, i.e., $\mathcal{W} = \emptyset$. Since $H = I$ and $c = 0$, we obtain $p^{(0)} = -(Hx^{(0)} + c) = -x^{(0)}$. The maximum steplength is given by

$$\alpha_{\max} = \min_{i: a_i^T p^{(0)} < 0} \frac{a_i^T x^{(0)} - b_i}{-a_i^T p^{(0)}} = \frac{2}{5},$$

where the minimium is attained for $i = 2$. Consequently, $\alpha^{(0)} = 2/5$ so that

$$x^{(1)} = x^{(0)} + \alpha^{(0)} p^{(0)} = \begin{pmatrix} 5 \\ 0 \end{pmatrix} + \frac{2}{5} \begin{pmatrix} -5 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 0 \end{pmatrix},$$

with $\mathcal{W} = \{2\}$. The solution to the corresponding equality-constrained quadratic progam is given by

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 0 \end{pmatrix} \begin{pmatrix} p_1^{(1)} \\ p_2^{(1)} \\ -\lambda_1^{(2)} \end{pmatrix} = - \begin{pmatrix} 3 \\ 0 \\ 0 \end{pmatrix}$$

We obtain

$$p^{(1)} = \begin{pmatrix} -\frac{12}{5} & \frac{6}{5} \end{pmatrix}^T .$$

The maximum steplength is given by

$$\alpha_{\max} = \min_{i: a_i^T p^{(0)} < 0} \frac{a_i^T x^{(0)} - b_i}{-a_i^T p^{(0)}} = \frac{5}{6},$$

where the minimium is attained for $i = 1$. Consequently, $\alpha^{(1)} = 5/6$ so that

$$x^{(2)} = x^{(1)} + \alpha^{(1)} p^{(1)} = \begin{pmatrix} 3 \\ 0 \end{pmatrix} + \frac{5}{6} \begin{pmatrix} -\frac{12}{5} \\ \frac{6}{5} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

with $\mathcal{W} = \{1, 2\}$. The solution to the corresponding equality-constrained quadratic progam is given by

$$\begin{pmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & 2 & 1 \\ 1 & 2 & 0 & 0 \\ 2 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} p_1^{(2)} \\ p_2^{(2)} \\ -\lambda_1^{(3)} \\ -\lambda_2^{(3)} \end{pmatrix} = - \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

We obtain

$$p^{(2)} = \begin{pmatrix} 0 & 0 \end{pmatrix}^T, \quad \lambda^{(3)} = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} \end{pmatrix}^T .$$

As $p^{(2)} = 0$ and $\lambda^{(3)} \geq 0$, the optimal solution has been found. Hence, $x^{(2)}$ is optimal.

3. If the problem is put on the form

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & g(x) \geq 0, \quad x \in I\!R^2, \end{array}$$

we obtain

$$\nabla f(x)^T = \begin{pmatrix} x_1 + x_2 + \frac{5}{2} & x_1 + x_2 - \frac{1}{2} \end{pmatrix}, \quad \nabla g(x)^T = \begin{pmatrix} x_2 & x_1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$\nabla^2_{xx}\mathcal{L}(x,\lambda) = \begin{pmatrix} 1 & 1-\lambda_1 \\ 1-\lambda_1 & 1 \end{pmatrix}.$$

With $x^{(0)} = (\frac{1}{2}\ 2)^T$ and $\lambda_1^{(0)} = 1$, the first QP-problem becomes

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}\begin{pmatrix} p_1 & p_2 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} p_1 \\ p_2 \end{pmatrix} + \begin{pmatrix} 5 & 2 \end{pmatrix}\begin{pmatrix} p_1 \\ p_2 \end{pmatrix} \\ \text{subject to} & \begin{pmatrix} 2 & \frac{1}{2} \\ 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} p_1 \\ p_2 \end{pmatrix} \geq \begin{pmatrix} 0 \\ -\frac{1}{2} \\ -2 \end{pmatrix}. \end{array}$$

The optimal solution of the QP-problem is given by the feasible point which is closest, in 2-norm, to $(-5\ \ -2)^T$, i.e., $p^{(0)} = (\frac{3}{17}\ \ -\frac{12}{17})^T$ with Lagrange multipliers $\lambda^{(1)} = (\frac{44}{17}\ \ 0\ 0)^T$. Thus, we have $\lambda^{(1)}$, and $x^{(1)}$ is given by $x^{(1)} = x^{(0)} + p^{(0)} = (\frac{23}{34}\ \frac{22}{17})^T$.

4. (See the course material.)

5. (a) By adding an additional variable z, we may rewrite (P) as the nonlinear program

$$(NLP) \qquad \begin{array}{ll} \text{minimize} & z \\ \text{subject to} & z - f_i(x) \geq 0, \quad i = 1,\dots,n, \\ & x \in I\!R^n, z \in I\!R. \end{array}$$

As f_i, $i = 1,\dots,n$, are convex on $I\!R^n$, (NLP) is a convex problem. Consequently, a local minimizer to (NLP) is also a global minimzier.
For a given positive μ, a barrier transformation of the constraints $z - f_i(x) \geq 0$, $i = 1,\dots,n$, gives the barrier function $B_\mu(z,x)$ on the form

$$B_\mu(z,x) = z - \mu \sum_{i=1}^{n} \ln(z - f_i(x)).$$

Minimizing $B_\mu(z,x)$ gives (NLP_μ), as required.

(b) The gradient of $B_\mu(z,x)$ is given by

$$\nabla B_\mu(z,x) = \begin{pmatrix} 1 - \mu \sum_{i=1}^{n} \frac{1}{z - f_i(x)} \\ \mu \sum_{i=1}^{n} \frac{1}{z - f_i(x)} \nabla f_i(x) \end{pmatrix}.$$

The first-order optimality conditions for minimizing $B_\mu(z,x)$ are given by $\nabla B_\mu(z,x) = 0$. By letting $\lambda_i = 1/(z - f_i(x))$, $i = 1,\dots,n$, we obtain the primal-dual nonlinear equations as

$$1 - \textstyle\sum_{i=1}^{n} \lambda_i = 0,$$

$$\begin{aligned} \textstyle\sum_{i=1}^n \nabla f_i(x)\lambda_i &= 0, \\ (z - f_i(x))\lambda_i &= \mu, \quad i = 1, \dots, n. \end{aligned}$$

As (NLP_μ) is a convex optimization problem, $B_\mu(z, x)$ is a convex function for z, x such that $z - f_i(x) > 0$, $i = 1, \dots, n$. To see this directly, we may form

$$\nabla^2 B_\mu(z, x) = \begin{pmatrix} 0 & 0 \\ 0 & \mu \sum_{i=1}^n \frac{1}{(z - f_i(x))^2} \nabla f_i(x) \nabla f_i(x)^T \end{pmatrix},$$

which is positive semidefinite for z, x such that $z - f_i(x) > 0$, $i = 1, \dots, n$. Consequently, a solution to $\nabla B_\mu(z, x) = 0$ corresponds to a global minimizer of (NLP_μ). Finally, the primal-dual nonlinear equations are equivalent to $\nabla B_\mu(z, x) = 0$.